Notes prepared by
Stanley Burris
March 13, 2001

Zermelo

Ernst Zermelo (1871–1953)
1904 - A proof that every set can be well ordered
1908 - Investigations in the foundations of set theory

In 1900 Hilbert had stated at the International Congress of Mathematicians that the question of whether every set could be well-ordered was one of the important problems of mathematics. Cantor had asserted this was true, and gave several faulty proofs. Then, in 1904, Zermelo published a proof that every set can be well-ordered, using the Axiom of Choice. The proof was regarded with suspicion by many. In 1908 he published a second proof, still using the Axiom of Choice. Shortly thereafter it was noted that the Axiom of Choice was actually equivalent to the Well-Ordering Principle (modulo the other axioms of set theory), and subsequently many equivalents were found, including Zorn's Lemma[1] and the linear ordering of sets (under embedding).

But more important for mathematics was the 1908 paper on general set theory. There he says:

> Set theory is that branch of mathematics whose task is to investigate the fundamental notions *number*, *order*, and *function* $\cdots$ to develop thereby the logical foundations of all of arithmetic and analysis $\cdots$. At present, however, the very existence of this discipline seems to be threatened by certain contradictions $\cdots$. In particular, in view of Russel's *antinomy* $\cdots$ it no longer seems admissible today to assign to an arbitrary logically definable notion a set, or class, as its extension $\cdots$. Under these circumstances there is at this point nothing left for us to do but to proceed in the opposite direction and, starting from set theory as it is historically given, to seek out the principles required for establishing the foundation of this mathematical discipline $\cdots$. Now in the present paper I intend to show how the entire theory created by Cantor and Dedekind can

[1] Originally proved by K. Kuratowski (1923) and R.L. Moore (1923); this was rediscovered by M. Zorn in 1935, and credited to him by Bourbaki.

> be reduced to a few definitions and seven principles, or axioms, $\cdots$. I have not yet even been able to prove rigorously that my axioms are consistent $\cdots$.

Zermelo starts off with a domain D of objects, among which are the sets. He includes $\approx$ and $\in$ in his language, and defines $\subseteq$. He says that an assertion φ is *definite* if

> the fundamental relations of the domain, by means of the axioms and universally valid laws of logic, determine whether it holds or not.

A formula $\varphi(x)$ is definite if for each x from the domain it is definite. Then the axioms of Zermelo are (slightly rephrased):

i. (Axiom of extension) If two sets have the same elements then they are equal.

ii. (Axiom of elementary sets) There is an empty set $\emptyset$; if a is in D then there is a set whose only member is a; if a, b are in D then there is a set whose only members are a and b.

iii. (Axiom of separation) Given a set A and a definite formula $\varphi(x)$ there is a subset B of A such that $x \in B$ iff $x \in A$ and $\varphi(x)$ holds.

iv. (Axiom of power set) To every set A there corresponds a set $P(A)$ whose members are precisely the subsets of A.

v. (Axiom of union) To every set A there corresponds a set $U(A)$ whose members are precisely those elements belonging to elements of A.

vi. (Axiom of choice) If A is a set of nonempty pairwise disjoint sets then there is a subset $C(A)$ of $U(A)$ which has exactly one member from each member of A.

vii. (Axiom of infinity)[2] There is at least one set I such that $\emptyset \in I$, and for each $a \in I$ we have $\{a\} \in I$

In the next few pages he defines $A \sim B$ (i.e., A and B are of the same cardinality),[3] derives Cantor's theorems that the cardinality of A is less than that of $P(A)$, and that every infinite set has a denumerably infinite subset.

[2] Later versions would use $a \cup \{a\} \in I$.

[3] Zermelo did not used ordered pairs. He starts with disjoint A and B and considers the set M in $P(A \cup B)$ consisting of doubletons with exactly one element from each of A and B. Then he looks for R in M which provide a 1–1 correspondence.

Let us use the traditional notation $\{a\}$ for *singleton*, and $\{a, b\}$ for *doubleton*. Let $\bigcup A$ be the *union* of the elements of A. One can define the basic set operations by

- $A \cap B \;=\; \{x \in A : x \in B\}$
- $\bigcap A \;=\; \{x \in \bigcup A : x \in a \text{ for all } a \in A\}$
- $A \cup B \;=\; \bigcup\{A, B\}$

We can let the set ω of *nonnegative integers*[4] be defined to be the smallest set A which contains $\emptyset$ and is closed under $x \in A \Rightarrow \{x\} \in A$. Then one has a successor operation $x' = \{x\}$ on ω which satisfies Peano's Axioms, so one can develop the number systems. Well, actually you need to have functions to do this. Using Kuratowski's definition of *ordered pair*, namely

$$(a, b) = \{\{a\}, \{a, b\}\},$$

one can prove (from Zermelo's axioms) that

$$(a, b) \approx (c, d) \qquad \Longleftrightarrow \qquad a \approx c \;\&\; b \approx d.$$

Then we can define
the *Cartesian product* of A and B:

$$A \times B = \{u \in P(P(A \cup B)) : x \in u \text{ iff } x \approx (a, b) \text{ for some } a \in A,\ b \in B\};$$

the set of *relations* between A and B:

$$Rel(A, B) = P(A \times B)$$

and the set of *functions* from A to B:

$$Func(A, B) = \{f \in Rel(A, B) : \forall a \in A\, \exists! b \in B \;\; (a, b) \in f\}.$$

With these definitions we can now translate Dedekind's development of the natural numbers, rationals, reals and complex numbers into Zermelo's set theory, and prove the basic properties about the operations $+, \times$; also we can carry out Cantor's study of sets, especially cardinals and ordinals. If one then wants to do analysis, for example integration on $[a, b]$, one takes the definite integral $\int_a^b dx$ as a certain element of $Func(\mathcal{F}, \mathbf{R})$, where $\mathcal{F}$ is a suitable subset of $Func([a, b], \mathbf{R})$; and then you can prove the fundamental theorem of calculus, etc., *all from Zermelo's seven axioms*.

However Zermelo's axioms had some obvious shortcomings. Skolem noted in *Some remarks on axiomatized set theory*, 1922, that improvements were needed, in particular

[4] Zermelo first looked at numbers in later articles.

- A definition of a *definite* property.
- An axiom to ensure some reasonably large sets.[5]

For the first he suggested one use first-order formulas, i.e., formulas made up from $\forall$, $\exists$, &, $\vee$, $\neg$, $\in$, $\approx$, and variables. (Skolem used the notation of Schröder.) Regarding the second he noted that if one has a model of set theory then let M_0 be the union of the $P^{(n)}(\{\emptyset\})$, the n-fold iterated power set applied to the set $\{\emptyset\}$, n an integer; and then let M be the union of the $P^{(n)}(M_0)$. This gives a rather small submodel, i.e., a small collection of sets that satisfies all of Zermelo's axioms. In particular the set $\{P^{(n)}(\omega) : n \in N\}$ is not a set in this model. To guarantee the existence of such interesting (not too large) sets he suggested the

- (Axiom of Replacement) if $\varphi(x, y)$ is a definite formula such that for every x there is at most one y making it true, then, for every set A there is a set B such that $y \in B$ holds iff there is an x in A such that $\varphi(x, y)$ holds.

Thus one can replace elements of A in the domain of φ with corresponding elements of B. Replacement is actually stronger than separation, for if one is given a set A and a definite property $\theta(x)$, define $\varphi(x, y)$ to be the definite property $y \approx x \,\&\, \theta(x)$. Then $\varphi(x, y)$ applied to A gives $\{x \in A : \theta(x)\}$.

The resulting set theory is called Zermelo-Fraenkel Choice (ZFC)[6].

Skolem pointed out that Zermelo's approach to set theory took us away from the natural and intuitive possibilities (like Frege's), and thus, as an artificial construction, carried a loss of status:

> Furthermore, it seems to be clear that, when founded in such an axiomatic way, set theory cannot remain a privileged logical theory; it is then placed on the same level as other axiomatic theories.

In 1917 Miramanoff noted the possibility of models with infinite descending chains $\cdots \in y \in x$. Such possibilities led von Neumann to formulate the Axiom of Regularity, namely if $x \not\approx \emptyset$ then for some $y \in x$ we have $x \cap y \approx \emptyset$. This axiom is not always used — it seems to have no application to mathematics, but it does make some proofs and definitions easier, e.g., that of an ordinal.

[5] Fraenkel published similar observations the same year.

[6] For a leisurely treatment, i.e., in the spirit of Zermelo's original paper, see Halmos' *Naive Set Theory*.

The treatment of ordinals evolved from Cantor's *abstraction* from well-ordered sets to *equivalence classes* of well-ordered sets to *representative* well-ordered sets. The last step was initiated by von Neumann in Transfinite numbers, 1923, and reached its modern brief form (assuming regularity) in the work of Raphael Robinson (1937), namely an *ordinal* is a transitive set (under $\in$), all of whose elements are also transitive sets.

ZFC requires an infinite number of first-order axioms. The open question of whether one could develop a set theory with a finite number of axioms was answered in the affirmative by J. von Neumann in 1925. Actually he used functions rather than sets as his primitive notion, and the current first-order version is due to the reworking in the late 1930's by (mainly) Bernays as well as Gödel, and called *von Neumann-Bernays-Gödel* set theory, abbreviated to NBG set theory.

Exercises

Problem 1 If R is a set of ordered pairs, show (using the axioms of ZFC) that the domain and the range of R are also sets.

Problem 2 Given ω and $+$ as sets, describe a set A and a first-order property $\varphi(x)$ such that the collection of integers Z is $\{x \in A : \varphi(x)\}$ (and thus it is a set by the axiom of separation). [Think of Z as sets of equivalence classes of ordered pairs of integers, where two pairs of integers are in the same class iff the first coordinate minus the second coordinate is the same in each case.]

Problem 3 [Kuratowski] Let us define (x, y) to be the set $\{\{x\}, \{x, y\}\}$. Use the axioms of Zermelo to prove that $(x, y) \approx (u, v) \iff x \approx u \,\&\, y \approx v$.

Could we use the definition $\{x, \{x, y\}\}$ and achieve the same?

Problem 4 Show that $x \notin x$ is a theorem in Z+(R)[7]

Problem 5 [R. Robinson] A set x is said to be *transitive* if $u \in v \in x$ implies $u \in x$. Suppose x and every element in x is a transitive set. Show that x is well-ordered by $\in$ (using Z+(R)).

References

[1] E. Zermelo, Investigations in the foundations of set theory I. 1908. [transl. in *From Frege to Gödel*, van Heijenoort, Harvard Univ. Press, 1971.]

[7] Zermelo's set theory with the axiom of regularity.